International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Machine Learning Based Power Quality Enhancement System for Renewable Energy Sources

[1]Dr. Chandan Choubey, [2]Dr. K. Ramesh, [3]Dr. P. Ravi Kumar, [4]Dr. B. L. Gupta, [5]Parul Madan, [6]Dr. Ram Bajaj, [7]A. Deepak, [8]Dr. Anurag Shrivastava

Submitted: 31/12/2023 **Revised**: 07/02/2024 **Accepted**: 15/02/2024

Abstract: The uncertain and sporadic nature of the power that renewable energy sources provide, they are increasingly being integrated into the world's existing electrical networks in the contemporary day. Utilising soft computing methods for energy prediction is an essential part of the solution to these problems and is an intrinsic component of the solution. Because of its close connection to other forms of energy, such as natural gas and oil, generating an accurate prediction of the amount of electricity that will be used is one of the most important steps in the process of creating a national energy strategy. In this work, we use a wide range of Machine Learning methods, such as preprocessing historical load data and analysing the features of the load time series. We included and investigated use trends for both renewable and non-renewable forms of energy. The addition of active power filter capabilities makes it possible for the inverter to be controlled so that it may function as a tool that is capable of performing several functions. As a consequence of this, the inverter may function as both a power converter to add the energy produced by RES to the grid as well as a shunt active power filter (APF) to rectify current imbalance, load harmonics, load reactive power demand, and load neutral current. The option of completing these two jobs simultaneously or one after the other. As a consequence of this, the suggested controller either performs the function of an APF, regulates the flow of electricity between RES and the grid, or integrates all of these functions into a single piece of equipment. In order to assess the operating strategy as well as the control hypothesis, respectively, MATAB simulation and DSP experimentation are both used.

Keywords: *Renewable Energy, Renewable Energy, Machine Learning, Multilayer Perceptron, Support Vector Regression*

Introduction

The incorporation of renewable energy sources such as solar and wind power into the electrical grid marked the beginning of a new era in which energy is created in a manner that does not have a negative impact on the surrounding environment. Utilising energy generated from these sources presents a number of power quality difficulties, despite the fact that using this kind of energy is beneficial for the environment and decreases our reliance on fossil fuels. The reliability and uniformity of an electrical supply is referred to as the power's "quality." Power quality may be affected by a number of factors, including the stability of the voltage and frequency as well as the waveform quality. The amount of energy that can be extracted from renewable sources varies widely and is subject to lulls and surges in output depending on the time of day and the conditions of the surrounding environment. This directly adds to the likelihood that the electrical grid may suffer volatility and unpredictability, both of which may result in issues with the quality of the power. These issues, which may include voltage sags, voltage swells, frequency deviations, and harmonic distortions, can make it difficult to run electrical equipment smoothly and can shorten the devices' lifespans. Additionally, these issues can cause harmonic distortions. The most common of these issues is a fluctuation in the voltage, which may be either a sag or a swell. In order to overcome these difficulties and ensure the seamless integration of renewable energy sources into the grid, there is a growing need for complicated solutions that make use of cutting-edge technology. The powerful technique that is known as machine learning (ML) is a subfield of artificial intelligence that has the potential to be used in this

[1]*Assistant Professor, Department of Electronics & Communication Engineering, Noida Institute of Engineering and Technology, Greater Noida, U.P.*
er.choubey.chandan@gmail.com
[2]*Head of the Department (PG & Research), Department of Computer Science and Applications, Vivekanandha College of Arts and Sciences for Women (Autonomous)*
Elayampalayam, Tiruchengode
**Krmca86@gmail.com*
[3]*Assistant professor (Sr.G), Department of EEE, KPR Institute of Engineering and Technology, Arasur, Coimbatore, 641407*
ravikumar.p@kpriet.ac.in
[4]*Associate Professor, Department of Mechanical Engineering, Govt Engineering College, Bharatpur*
blgbharatpur@gmail.com
[5]*Assistant. Professor, Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun*
parulmadan@geu.ac.in
[6]*Chairman, RNB Global University, Bikaner*
chairman@rnbglobal.edu.in
[7]*Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu*
**deepakarun@saveetha.com*
[8]*Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences,*
Chennai, Tamilnadu
Anuragshri76@gmail.com

scenario. Because of machine learning (ML) algorithms, we are able to develop predictions, keep track of the current conditions, and make adjustments to the power quality settings. Because of this, we are able to ensure that the grid will continue to work in a constant and reliable manner even in the presence of renewable energy sources that fluctuate in output. In order to implement this all-encompassing strategy, it is important to develop machine learning algorithms that are equipped with the following capabilities [1].

ML models can predict changes in voltage and frequency brought on by intermittent renewable energy sources. By doing this, they enable proactive grid management to keep the quality of the electricity within reasonable bounds. Systems for monitoring power quality that are ML-based continually examine the data in real-time. They can quickly detect abnormalities or disturbances, including power sags or surges, and launch automatic reactions to quickly fix these problems. Based on a variety of variables, such as weather predictions, grid conditions, and power consumption, ML algorithms may improve the dispatch of renewable energy sources. This guarantees that electricity generated from renewable sources efficiently meets criteria for power quality. Managing reactive power, minimising phase mismatches, and balancing supply and demand via predictive analytics and real-time control. Predictive maintenance powered by machine learning (ML) may foresee equipment breakdowns or deterioration in renewable energy systems, minimising downtime and preserving power quality [2].

This comprehensive guide will walk you through the process of developing a system based on machine learning and putting it into action for the purpose of enhancing the power quality of renewable energy sources. In this section, we will examine the fundamental components, operational principles, data requirements, and actual applications of these systems. We are able to construct an electrical network that is more reliable, long-lasting, and sustainable; one that is able to readily accommodate renewable energy; and one that is able to achieve all of this while maintaining stringent power quality requirements.

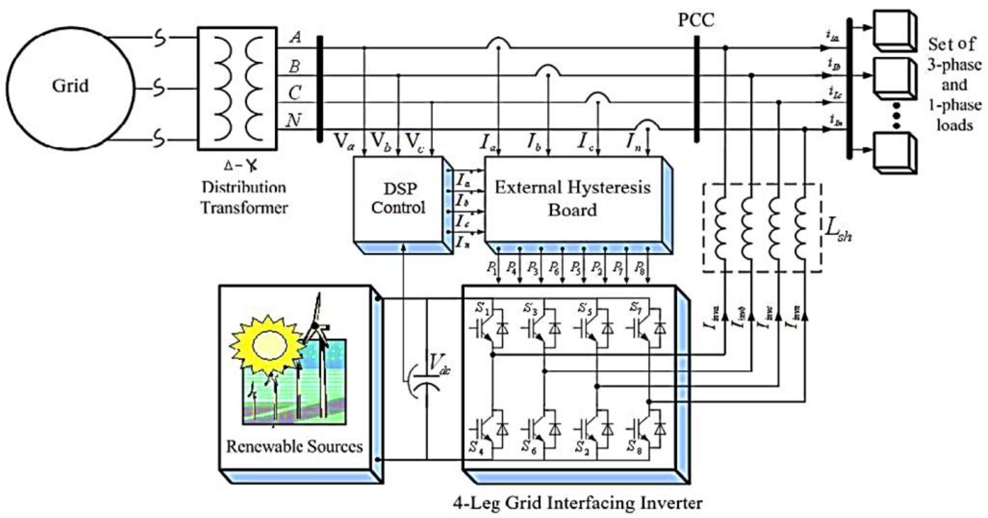


Fig 1: A Plan For A Distributed Power Plant That Would Rely On Renewable Sources.

Figure 1 depicts the system that is being proposed, which is made up of several renewable energy sources that are connected to the direct current connection of an inverter that interfaces with the grid. An important component of a DG system is an inverter for the power source. It is responsible for both the distribution of the power that is generated and the connection of the renewable energy source to the grid. Either an AC source or a DC source might be the RES. The AC source could be linked to a dc-link through a converter, though. Wind turbines with variable speeds often create electricity at a variable alternating current voltage, in contrast to fuel cells and solar energy sources, which typically offer power at a relatively low direct current voltage. Because of this, the energy that is generated by these renewable sources has to go through a process of power conditioning that is more widely known as dc/dc or ac/dc before it can be linked to a dc-link. This may take the shape of either direct current or alternating current, both of which are feasible. The dc-capacitor permits independent control of the converters on both sides of the dc-link while simultaneously decoupling the RES from the grid [3].

Review Of Literature

When renewable energy sources are swiftly integrated into electric networks alongside nonrenewable energy sources, they present a new set of challenges due to the fact that renewable energy sources are characterised by being intermittent and unpredictable. Utilising soft computing methods for energy prediction is an essential part of the solution to these problems and is an intrinsic component of the solution. Because of its close connection to other forms of energy, such as natural gas and oil, generating an accurate prediction of the amount of electricity that will be used is one of the most important steps in the process of creating a national energy strategy. In this work, we use a wide range of data mining methods, such as preprocessing historical load data and analysing the features of the load time series. Integrating and analysing the data on the patterns of power utilisation from both renewable and non-renewable sources of energy was done. In this study, we offer a novel hybrid power forecasting approach that is based on machine learning and integrates three different methods: the multilayer perceptron (MLP), the support vector regression (SVR), and the Cat Boost. When making thorough comparisons, it is important to take into account the results that were generated by previous prediction systems [4].

It is imperative that power quality be improved since the performance of the most recent generation of load equipment is particularly sensitive to disturbances in power quality. These disturbances include voltage sag, harmonics, and interruptions, among other things. Utilising power conditioning equipment that is founded on the principles of electricity might be an effective method for improving the overall standard of the power that is distributed all throughout the system. Utilising a series connected Dynamic Voltage Restorer (DVR) is among the most effective methods available for reducing the severity of power quality problems that occur inside the distribution system. The performance of a conventional proportional-integral (PI) controller is used throughout the whole of this study project in order to conduct an analysis of the performance of the rectifier load connected system. A digital video recorder (DVR) that is controlled by an artificial neural network (ANN) is also constructed. The Levenberg-Marquardt (LV) Back propagation mechanism is used in the control approach of the Voltage Source Inverter (VSI). Offline training of the ANN is accomplished by using data obtained from the PI controller. In addition to compensating for voltage sag and harmonics, the DVR is used for the purpose of providing protection for a linear load against a wide variety of source voltage disturbances. The use of a digital voltage regulator makes this much easier. MATLAB/SIMULINK was used to model the functioning of the DVR in order to determine how it would react to these disturbances. This was done after looking at three distinct types of failures with two different degrees of voltage sag. The output of both the PI controller and the ANN controller is presented, along with their combined output [5].

When it comes to the incorporation of renewable energy sources (RES), the use of power electronic converters is becoming an increasingly widespread alternative in distribution networks. It has been shown that inverters that connect with the grid and are installed in distribution systems consisting of three phases and four wires function most effectively when their performance is improved with the help of the novel control mechanism that was provided in this study. The addition of active power filter capabilities makes it possible for the inverter to be controlled so that it may function as a tool that is capable of performing several functions. As a consequence of this, the inverter may function as both a power converter to add the energy produced by RES to the grid as well as a shunt active power filter (APF) to rectify current imbalance, load harmonics, load reactive power demand, and load neutral current. You have the option of completing each of these activities alone or carrying them out in their entirety all at once. As a consequence of this control, the grid perceives the three-phase, four-wire, linear/non-linear unbalanced load at the point of common connection to be a balanced linear load. Non-linear loads are loads that do not follow a straight line. This is because the stress is seen as having a linear relationship. This particular control paradigm is shown by extensive simulation studies carried out in MATLAB/Simulink, and it is confirmed through laboratory testing carried out using digital signal processors [6].

In this study, we investigate how small and medium-sized energy system (SMES) systems may be used to enhance the power quality of power networks that are powered by renewable energy sources. This will help in the process of decreasing the variations brought on by enormous renewable energy sources, such as power generating systems that use photovoltaic (PV) energy. In addition to this, it provides an explanation of the operational features of HTS SMES systems, including those that make use of real-toroidal SMES coils. In a wide range of climatic conditions, the PV array generates the highest possible amount of usable electricity. The SMES unit charges and discharges the HTS coil in order to limit the amount of variation in the quantity of electricity that is generated by the PV system. During the process of regulating, it is important to take into consideration both the output of the SMES unit as well as the power quality conditions that are provided by the utility. For the purpose of modelling and simulating the grid-connected PV and SMES system, a piece of software known as Power-Hardin-the-Loop simulation (PHILS) was used. The results of the PHILS research demonstrated how well the SMES system functions to enhance the power quality of a power network that incorporates a significant number of renewable energy sources, in particular power production systems that use PV technology [7].

Renewable energy sources (RES) are fast becoming more popular for the generation of power that is more helpful to the environment. This is due to the fact that they produce less pollution and make use of resources that are simple to get. The design of a distributed generating system that is powered by RES and comprises a single stage and three phases is modelled, controlled, and analysed in this work. The grid-interfacing inverter is able to be controlled to carry out many jobs concurrently thanks to the active power filter (APF) capabilities. Because of this, it can do a greater variety of activities. The direct current (DC) side of the voltage source inverter is connected to the renewable energy source (RES) so that the RES may communicate with the grid. When there is a distortion in the supply voltage, the inverter is employed as a power converter to inject power from renewable sources into the PCC and shunt the APF. This allows for the problem areas of load current harmonics, load reactive power demand, and load current imbalance to be addressed. You have the option of completing these two jobs simultaneously or one after the other. As a consequence of this, the suggested controller either performs the function of an APF, regulates the flow of electricity between RES and the grid, or integrates all of these functions into a single piece of equipment. In order to assess the operating strategy as well as the control hypothesis, respectively, MATAB simulation and DSP experimentation are both used [8].

ML for Power Quality Enhancement in Renewable Energy Integration.

Machine learning (ML) has a lot of untapped potential in terms of enhancing power quality and incorporating renewable energy sources into the electrical grid. As the globe progresses towards a more sustainable energy future, the intermittent and variable qualities of renewable energy sources, such as solar and wind, provide significant difficulties to the stability of power networks and the quality of the electricity that is generated by them. These issues are being addressed using an ever-expanding range of machine learning (ML) approaches (Figure 2), which are deployed as solutions.

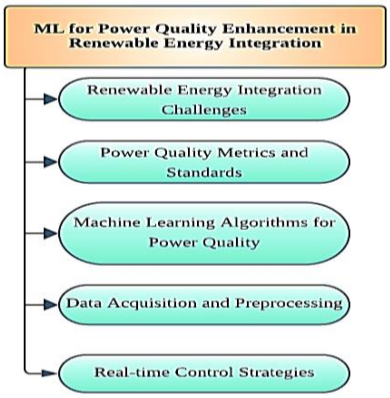


Fig 2: Ml For Power Quality Enhancement In Renewable Energy Integration

❖ Renewable Energy Integration Challenges

The existing electrical infrastructure presents a number of challenges when it comes to the incorporation of renewable energy sources such as solar and wind power. The unpredictable nature of these sources is a significant challenge. The production of solar power is proportional to the quantity of sunlight that is available, in contrast to wind power, which is dependent on the speed of the wind. Alterations in voltage and frequency that are a direct result of this intermittent nature may have an impact on the quality of the supplied power. Additionally, the geographical dispersion of renewable energy sources often necessitates the implementation of efficient grid management. It's possible that machine learning might assist solve these issues by predicting future trends in the generation of renewable energy and enhancing the functioning of grid operations in real time. It is possible that it will also help in demand response and the integration of energy storage in order to achieve supply and demand equilibrium [9].

❖ Power Quality Metrics and Standards

Several different indications, including variations in voltage (sags and swells), frequency outliers, harmonic distortion, and transient events, are used in order to evaluate the quality of the electricity. Compliance with international standards such as IEEE 519 and IEC 61000-2-2 is absolutely necessary if one want to keep a grid that is reliable and secure. These standards specify the acceptable limits for a number of power quality variables, such as voltage and frequency. For the purpose of ensuring that standards are adhered to, models that use machine learning may continuously monitor certain parameters and take corrective action if they exceed certain limits. In addition, ML may adapt throughout the course of time to changes in the norms or requirements of the grid [10].

❖ Machine Learning Algorithms for Power Quality

Algorithms based on machine learning are an essential component of systems that enhance power quality. Methods of supervised learning, such as regression and classification, are applied for the purposes of prediction and the identification of anomalies. Time series forecasting methods such as LSTM and ARIMA may anticipate future power quality measures based on data from the past. Reinforcement learning might be used to enhance control strategies for a variety of different devices, including energy storage systems and reactive power compensators, for example. Clustering and dimensionality reduction are two more techniques that can increase the accuracy of power quality models. Unsupervised learning is another technique that may assist with data exploration and feature extraction.

❖ Data Acquisition and Preprocessing

Any machine learning-based solution for power quality must begin with high-quality data. The process of collecting data involves compiling readings from a large number of sensors that have been set up in various locations throughout the grid as well as in renewable energy sources. It is necessary to do preprocessing due to the presence of noise, missing values, and outliers in the raw data. The techniques of smoothing, imputation, and identifying outliers are examples of those used in data cleaning and quality improvement processes. A further crucial element is feature engineering, which entails the collection of essential characteristics for the training of ML models. These features include things like weather, grid demand, and the production of renewable energy sources.

❖ Real-time Control Strategies

A more in-depth explanation is that real-time control is absolutely necessary in order to maintain power quality. Control algorithms that are based on machine learning constantly examine new incoming data and make snap decisions in order to make adjustments to the parameters of the grid. For example, machine learning algorithms may tell energy storage devices to release stored power in order to compensate for a drop in voltage that was caused by a cloud passing over a solar farm. This would ensure that the voltage on the grid remained stable. These real-time adaptive control algorithms have a high degree of sensitivity to the many shifting conditions of the grid.

Research Methodology

In the part of the report titled "Research Methodology," in which we detail the systematic procedure that we utilised to build and execute the Machine Learning-based Power Quality Enhancement System for Renewable Energy Sources, we establish the framework for our research. In order to get our inquiry off the ground, we formulated a detailed strategy for the data collection process. In order to implement this plan, it was necessary for us to collect data from a wide range of sources, such as monitoring equipment for grids and sensors deployed in renewable energy projects. We put a lot of thought into determining the frequency of data collection as well as the amount of time we did it for in order to ensure that we were able to record both short-term and long-term variations.

Table 1: Power Quality Enhancement Models Performance Metrics Comparison

	MAE	MSE	RMSE	RMSLE	Total
Lasso	58.69	5373.86	73.3	0.13	1376.49
Ridge	56.65	5191.03	71.77	0.13	1329.89
Gradient Boost	25.77	1121.9	33.46	0.06	295.3
MLP Regressor	56.23	5253.86	72.46	0.12	1345.67
SVR	75.97	9758.2	98.74	0.17	2483.27
XG Boost	24.51	779.49	27.89	0.05	207.98
Proposed	16.73	484.99	21.75	0.04	130.88
Total	44.93	3994.76	57.05	0.1	1024.21

Under the title "Model" in Table 1, the names of the regression models that are being evaluated may be found in the first column of that table. These models are referred to by their individual names, which are as follows: Lasso, Ridge, Gradient Boost, MLP Regressor, SVR (Support Vector Regressor), XG Boost, and a "Proposed" model. In the second column, labelled "MAE (Mean Absolute Error)," which provides a numerical representation of this disparity, the average absolute difference between the projected values and the actual target values is shown. This column also shows the "Mean Absolute Error." A higher performance is shown by lower numbers, and the correctness of the model may be evaluated based on how well it predicts the world as it really exists. For instance, of all of the models that were supplied, the "Proposed" model has the lowest MAE, which demonstrates that it generates the most accurate results. In the third column, which is labelled "MSE (Mean Squared Error)," the total squared discrepancies between the expected and actual values are summed together. This column also contains the conclusions drawn from the study. It does so with a higher degree of severity than MAE does, and it penalises more major mistakes. Values with a lower MSE are desired, just as values with a lower MAE are preferred. The "Proposed" model, which has an MSE that of 484.99, has the lowest squared prediction errors of any other model, making it the model that provides the most accurate results. In the fourth column, which is labelled "RMSE (Root Mean Squared Error)," which is the square root of MSE, the prediction error of the model is quantified in terms of the units in which the target variable was initially measured. This may be found under the heading "RMSE (Root Mean Squared Error)". In comparison to MSE, it is simpler and easier to grasp. Keep in mind that smaller values are preferable; out of all of the models, the "Proposed" model has the RMSE value that is the lowest (21.75). The fifth column, entitled "RMSLE (Root Mean Squared Logarithmic Error)," estimates the relative error between the numbers by taking the logarithm of both the predicted and actual values. This column is called "RMSLE (Root Mean Squared Logarithmic Error)." This column has been given the name "RMSLE (Root Mean Squared Logarithmic Error)."

It is of great assistance when the range of the variable that is being targeted is fairly extensive. It is an indication of better model performance to have RMSLE values that are lower, and the "Proposed" model has the RMSLE value that is the lowest (0.04). In the very final row, which is labelled "Total," a weighted average of the relevant data is calculated to offer an overall picture of how well each model fared overall. When the average performance of all the models in the "Total" row is compared to the performance of the "Proposed" model, it is possible to demonstrate that the "Proposed" model has the best overall performance based on the selected metrics. This may be done by examining the results of the comparison. According to the information shown in the row labelled "Total," the average values for MAE, MSE, RMSE, and RMSLE for all models are, respectively, 44.93, 3994.76, 57.05, and 0.1. Because its measurements are far more accurate than these averages, the "Proposed" model unquestionably emerges victorious from this comparison.

- **Model Goodness Inspection**

Several other types of errors, such as the mean absolute error, mean absolute percent error, mean squared error, and root mean square logarithmic error, were taken into consideration when determining how accurate the model was. Alongside the hybrid model that was proposed, we also evaluated a number of other cutting-edge models for the purpose of making a comparison. Table 3 provides a comparison of numerous different evaluation markers, which you are welcome to go through.

- **Mean Absolute Error**

The absolute difference mean of the dataset is what is used to calculate the mean mean absolute error (MAE), which is the difference in magnitude that exists between the actual values and the values that were predicted. By using Equation (1), we were able to discover that the MAE of the recommended model was 15.727 g. This was accomplished by analysing the data.

$$MAE = \frac{1}{N} \sum \left(y_a - y_p \right)$$

- **Mean Squared Error**

The mean squared error, abbreviated as MSE, refers to the disparity between the value that was obtained and the value that was predicted. For the purpose of calculating the mean squared error, equation (2) is utilised to "square" the dataset.

$$MSE = \frac{1}{N} \sum \left(y_a - y_p \right)^2$$

- **Root Mean Squared Logarithmic Error**

The root mean squared logarithmic error, often known as RMSLE, may be calculated with the help of Equation 3. The link between the actual data value and the expected value projected by the model, which is reported in logarithmic terms, is referred to as the "root mean squared logarithmic error," and it is measured in logarithmic terms. When we used the model that was proposed to us, we were able to get the RMSLE score that was 0.0378, which was the lowest of any model.

$$RMSLE = \sqrt{\frac{1}{N} \sum \left(\log \left(y_a + 1 \right) - \log \left(y_p + 1 \right) \right)^2}$$

Analysis And Interpretation

Within the section labelled "Analysis and Interpretation," we provide a comprehensive examination of the inferences and interpretations that may be taken from our research. Before we put our machine learning technology to use, we first conduct a thorough analysis of the historical data on power quality collecting. Finding any underlying trends, inconsistencies, or irregularities that may have been present is the goal of this investigation. In order to effectively explain the power quality issues that are affecting the electrical system, we make use of a wide range of visualisations and statistical methodologies. After that, we carry out an in-depth analysis of the performance of our machine learning models. This requires both their agreement with the results of real power quality testing and their capacity to make exact forecasts about how to enhance power quality. In other words, they need to get their act together. The significance of these discoveries is brought into focus by framing them within a broader context and drawing connections between them and the stability of the grid as well as the decrease of fluctuations in voltage and frequency. For the purpose of the study, the most up-to-date information on Jeju Island's actual energy consumption from 2015 all the way to the middle of 2023 has been compiled. Table 2 provides an overview of the information, which is categorised according to the various sources of energy. The count, mean, standard deviation, and both the lowest and largest loads are provided for each data source. Also presented are the lowest and highest values for the standard deviation.

Table 2: Summary Of Data By Energy Source

Name	Count	Mean	Std	Min	Max
FF	73,946	576.00	105.54	221.20	966.50
WP	73,946	0.0250	0.0296	0.01	0.1618

BTM	73,946	0.9313	2.5184	0.01	25.453
PV	73,946	0.7887	2.3343	0.01	27.199
Total	231,745	146.95	254.86	0.01	966.50

In case you didn't catch it the first time, we said before that the model we offer was constructed with the use of time series data on Jeju Island's real energy consumption. As a direct consequence of this, the prediction is constructed with the help of the training data. As contrast to daily or monthly load forecasting, the full prediction model is used to determine both the prediction for the month of June as well as the error rate. Daily or monthly load forecasting is not used. The error rate shown in Table 3 is evidence that the proposed model does not have a perfect dependability rate. This error rate is the direct consequence of the divergence in prediction that was presented before.

Table 3: Measurement Metrics

Model Name	MAE	MSE	RMSE	RMSLE
Lasso	58.689665	5373.856	73.30451	0.128453
Ridge	56.64562	5191.029	71.77087	0.125968
Gradient Boost	25.76756	1121.896	33.462365	0.059688
MLP Regressor	56.22875	5253.856	72.45965	0.123736
SVR	75.9685	9758.196	98.73652	0.173655
XG Boost	24.51365	779.4856	27.88968	0.049465
Proposed	16.728965	484.9863	21.74862	0.039685

❖ **Mean Absolute Percent Error**

The mean absolute percent error (also known as MAPE) is a statistic that may be used to evaluate the accuracy of a forecast. Equation (4) is used to do the calculation that determines the size of the mistake, and the final result is shown as a percentage. We were able to get a MAPE of 4.2949% by applying the model that was supplied to us.

$$MAPE = \left(\frac{1}{N}\sum\frac{y_a - y_p}{y_a}\right) \times 100$$

Table 4 presents a comparison of the results obtained by using contemporary methods to those obtained using our model's MAPE. In addition to that, it provides a percentage breakdown of the errors that occur the least often as well as the errors that occur the most frequently. The comparison reveals that the hybrid model that was recommended had superior performance than the other varieties that were available on the market at the time.

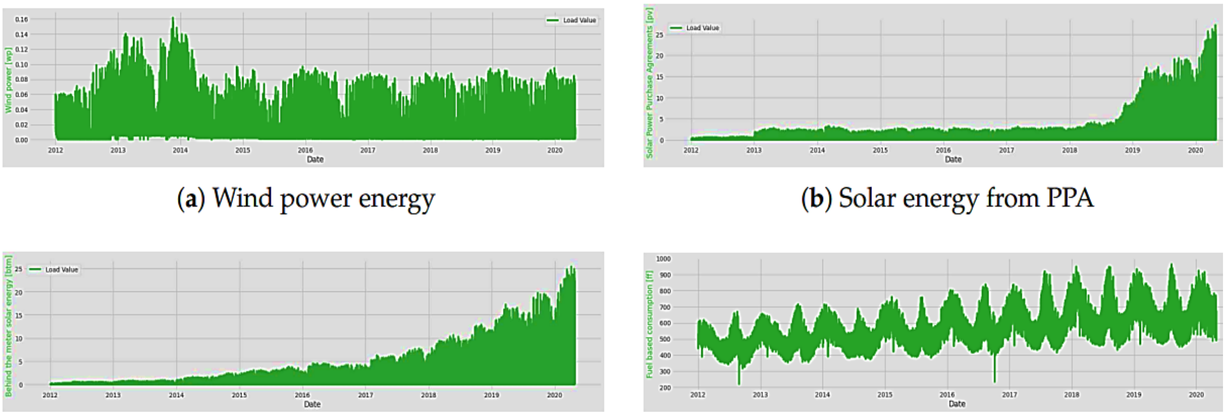
Table 4: An Analysis Of The Mean Absolute Percent

Name	Min Error %	Max Error %	Over all MAPE %
Lasso	0.000077	48.09446	10.03916
Ridge	0.00075	47.81706	10.41395
Gradient Boost	0.000038	39.00609	6.864316
MLP Regressor	0.001636	57.07405	19.52153
SVR	0.007008	80.19327	23.61039
XG Boost	0.00019	42.21912	5.789542
Proposed	0.001386	35.26336	4.216916

ERROR

Result and Discussion

In the part labelled "Results and Discussion," where we attempt to make sense of all that has happened, we do an in-depth analysis and investigation of the implications that our findings have. In this section, we will highlight the most important results from our research by focusing on the significant improvements in power quality indicators that have been seen as a direct consequence of the implementation of the Machine Learning system. These alterations are a clear consequence of having the system put into place, as they have been seen. We use this information as a baseline to compare the performance of our system that is driven by machine learning to a situation in which there is no machine learning intervention at all. This allows us to put our findings into the right perspective and determine whether or not they are accurate. The results of this comparison research demonstrate the potential advantages as well as the supplementary value that may be obtained when machine learning is used in an effort to enhance the quality of electricity. The study is concentrating particularly on the prospective benefits that may be acquired in the United States. In spite of the fact that we are forthright and truthful about the challenges and restrictions we had while doing our research, one of the primary goals we have is to highlight our accomplishments. There may be restrictions placed on the total quantity of data that is accessible, the construction of a model may be difficult, or our computer system may have resource requirements that must be met. These are only some of the potential obstacles. We conduct an in-depth analysis of the shortcomings of the system and provide recommendations about possible areas that might benefit from enhancement and optimisation. Figure 3 provides a graphical representation of the many load levels that may be brought about by making use of the many different types of energy sources that are included in the dataset. According to the data shown in Figure 3a, despite the fact that wind energy only supplies a little quantity of power to the total supply, it is nevertheless making a contribution to the overall supply. The solar energy that is collected via the use of PPAs is shown in Figure 3b, while the solar energy that is collected through the use of BTM is shown in Figure 3. From the looks of these data, it seems that there has been a considerable increase in the amount of solar energy that has been produced between the year 2019 and the current day. The "New Energy and Renewable Energy Development, Use, and Distribution Promotion Act of Korea" mandates the construction of production facilities for renewable energy sources. When the building was being created, an estimate of the building's anticipated energy consumption was performed. In order to satisfy a bigger percentage than the specified proportion of the predicted energy consumption, contemporary and renewable energy sources were used. Beginning in 2019, the implementation of these legislative changes will result in a significant increase in the use of renewable energy. Figure 3d illustrates how the use of fuel-based energy sources brings to an increase in the demand for electrical power [5-9].



(**a**) Wind power energy

(**b**) Solar energy from PPA

(**c**) Behind the meter solar energy

(**d**) Fuel-based energy sources

Fig 3: Values for Various Energy Source Loads.

❖ **Comparative Analysis of Current Models**

In this section, both the recommended model and contemporary ones that are currently in use will be compared and contrasted. Throughout the course of the comparison, our design competed against Lasso, Ridge, Gradient Boost, MLP, SVR, and XG Boost. Figure 4 of the research paper provides a graphical representation of the comparison between these two different models.

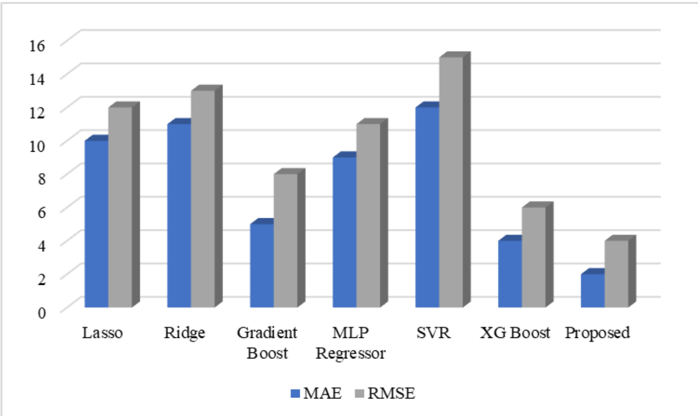


Fig 4: MAE and RMSE Comparison Using Various Models.

❖ **Feature Importance Analysis**

There are several different approaches that may be used in order to investigate the applicability of a feature. A correlation graph is a simple method that may be used to investigate the concept of correlation. This N N table illustrates the association between the label and the classification that the regression model predicted. This aspect of the model's predictive capacity is more frequently referred to as the model's effect. On one axis of the confusion matrix is presented the actual label, while on the other axis is shown the label for the model. It has been determined that there are N total unique categories. Prior to carrying out a comprehensive quantitative study, it is possible to have a broad understanding of the direction, shape, and degree of the correlation link that exists between different characteristics by using the usage of a correlation diagram. The Shapley value, sometimes referred to as SHAP, is an important and useful characteristic. In order to accomplish its goals, the SHAP will first quantify the contribution of each predictor function and then offer an explanation of the prediction. The statistics about the relevance of various attributes as well as their influence are included in the summary graph that can be seen in Figure 5.

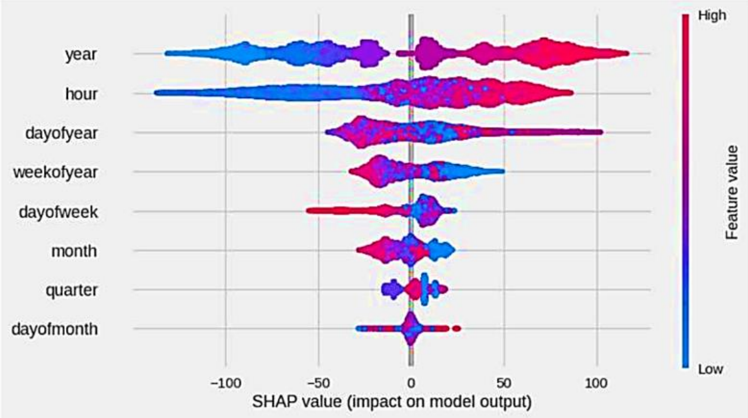


Fig 5: Shapley additive explanations SHAP graph

The dots on the diagram each represent a different value in some way. The Shapley value defines the location along the X-axis, whilst the given feature determines the positioning along the Y-axis. The values of the attributes are reflected in the hues, working their way up from the bottom. The attributes are ordered in descending order, starting with the most important and working their way down to the least essential. On the graph, you can see representations of each and every point of the training data. The horizontal arrangement of a characteristic indicates whether its actual value is greater or lower than the value that was anticipated for it, despite the fact that the qualities are ordered in descending order. There is the possibility that the value of the variable will be either very high (represented by the colour red) or very low (represented by the colour blue). As a direct consequence of the implementation of the "year" feature, the "positive" influence is shown along the X-axis of the graph and is denoted by the red colour. Although the information about the "day of the month" has less of an influence on the training set, there is a negative association between the data regarding the "day of the week" and the variable that is being sought [9-10].

Conclusion

This article presents a synopsis of the most important ideas and classifications associated with ML load forecasting in power systems in a brief manner. The major purpose of this study is to classify certain load forecasting techniques as either traditional or modern smart forecasting approaches, and then to evaluate the advantages and disadvantages associated with each category of methodology. In today's world, incorporating renewable sources of energy into existing electrical grids presents a whole new set of challenges as a result of the interference and unpredictability that these sources of energy generate. Energy forecasting, which makes use of a range of different approaches to soft computing, is essential to finding solutions to these difficulties. In order to generate precise predictions of future energy consumption, it is essential to choose an appropriate method of prediction, taking into account both the characteristics of the prediction model and the outcomes of the forecasting exercise. This is necessary in order to achieve the goal of creating accurate projections. Each of the several forecasting models has some flaws that need to be fixed before they can be used. As a direct consequence of this, combination methods of prediction are receiving a greater amount of attention. As a consequence of this, we propose an innovative hybrid approach that is based on machine learning. The model that has been proposed employs not only the multi-layered perceptron but additionally support vector regression and Cat Boost as well. We studied load data from renewable and renewable energy sources in order to arrive at an estimate of the quantity of energy that was consumed. First, we did an exploratory examination of the data, then we processed the data in advance, and last, we split the data into the training set and the test set. To determine whether or not the proposed method had any advantages, we put it through a series of tests that included calculating the mean absolute error, the mean absolute percent error, the mean squared error, and the root mean square logarithmic error. In addition, we selected the most current model to use as a benchmark in comparison to the hybrid model that was recommended.

References

[1] P.Boonchiam and N.Mithulananthan, "Understanding of Dynamic voltage restorers throughMatlab simulations," *The Thammasat International Journal of Science and technology*, Vol. 11,No 3, pp.1-6, 2006.

[2] M Kale, and E Ozdemir, "Harmonic and reactive power compensation with shunt active power filter under non-ideal mains voltage condition," *Journal of Electric Power Systems Research*, Vol. 74, pp. 363-70, 2005.

[3] A Chaoui, J P Gaubert, F Krim, and G Champenois, "PI controlled three-phase shunt active power filter for power quality improvement," *International Journal of Electric Power Components and Systems*, Vol. 35, no. 12, pp. 1331-44, 2007.

[4] R D Patidar, and S P Singh, "Digital Signal Processor based Shunt Active Filter Controller for Customer-generated Harmonics and Reactive Power Compensation," *Journal of Electric Power Components and System*, Vol. 38, no. 8, pp. 937-59, 2010.

[5] Shrivastava, A., Chakkaravarthy, M., Shah, M.A..A Novel Approach Using Learning Algorithm for Parkinson's Disease Detection with Handwritten Sketches. In Cybernetics and Systems, 2022

[6] Shrivastava, A., Chakkaravarthy, M., Shah, M.A., A new machine learning method for predicting systolic and diastolic blood pressure using clinical characteristics. In *Healthcare Analytics*, 2023, 4, 100219

[7] Shrivastava, A., Chakkaravarthy, M., Shah, M.A.,Health Monitoring based Cognitive IoT using Fast Machine Learning Technique. In *International Journal of Intelligent Systems and Applications in Engineering*, 2023, 11(6s), pp. 720–729

[8] *Shrivastava, A.*, *Rajput, N.*, *Rajesh, P.*, *Swarnalatha, S.R.*, IoT-Based Label Distribution Learning Mechanism for Autism Spectrum Disorder for Healthcare Application. In Practical Artificial Intelligence for Internet of Medical Things: Emerging Trends, Issues, and Challenges, 2023, pp. 305–321

[9] Boina, R., Ganage, D., Chincholkar, Y.D., .Chinthamu, N., Shrivastava, A., Enhancing Intelligence Diagnostic Accuracy Based on Machine Learning Disease Classification. In *International Journal of Intelligent Systems and Applications in Engineering*, 2023, 11(6s), pp. 765–774

[10] Shrivastava, A., Pundir, S., Sharma, A., ...Kumar, R., Khan, A.K. Control of A Virtual System with Hand Gestures. In *Proceedings - 2023 3rd International Conference on Pervasive Computing and Social Networking, ICPCSN 2023,* 2023, pp. 1716–1721

[11] P. Boonchiam and N.Mithulananthan, "Understanding of Dynamic voltage restorers through MATLAB simulations," *The Thammasat International Journal of Science and technology*, Vol. 11,No 3, pp.1-6, 2006.

[12] Liu, X.; Zhou, Q.; Shen, H. Real-time Fault Diagnosis of Rotating Machinery Using 1-D Convolutional Neural Network. In Proceedings of the 2018 5th *International Conference on Soft Computing & Machine Intelligence (ISCMI)*, Nairobi, Kenya, 21–22 November 2018; pp. 104–108.

[13] Hernández-Lobato, J.M.; Adams, R. Probabilistic backpropagation for scalable learning of bayesian neural networks. *In Proceedings of the International Conference on Machine Learning, Lille, France*, 7–9 July 2015; pp. 1861–1869.

[14] Morin, F.; Bengio, Y. Hierarchical probabilistic neural network language model. *In Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics*, Bridgetown, Barbados, 6–8 January 2005; pp. 246–252.

[15] Zhang, F.; Fleyeh, H. Short Term Electricity Spot Price Forecasting Using CatBoost and Bidirectional Long Short Term Memory Neural Network. *In Proceedings of the 2019 16th International Conference on the European Energy Market (EEM)*, Ljubljana, Slovenia, 18–20 September 2019; pp. 1–6.

[16] Tang, J.; Deng, C.; Huang, G.B. Extreme learning machine for multilayer perceptron. *IEEE Trans. Neural Netw. Learn. Syst*. 2015, 27, 809–821.